How Well are Domain and Upper Ontologies Connected?

Ivan Opalički, Sandra Lovrenčić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{ivan.opalicki, sandra.lovrencic}@foi.hr

Abstract. *There is a substantial number of various proposals for possible applications of upper ontologies. Several comparisons of chosen upper ontologies have also been made and an idea of one unified upper ontology is still present. This paper describes an initial research about connections of domain ontologies with upper ones through relations between root classes of domain ontologies and upper ontology concepts. Analysis results and potential for further research are also presented.*

Keywords. upper ontology, top-level ontology, foundation ontology, domain ontology, ontology development

1 Introduction

There are several well-known upper ontologies that describe the most general concepts and their relations. Even from the beginning of their development, there has been an idea that only one unified upper ontology should be a final result of this research [4].

Over years, practical use of upper ontologies has been suggested for various purposes, some of them being development of interoperating information systems [2], fostering of interoperability between Simple Knowledge Organization System (SKOS) concept schemes [12] and Linked Open Data (LOD) ontology alignment [6]. They should also be a starting point for development of ontologies on more detailed level of description (specific domain ontologies) and recent research has in fact proven that domain ontologies can benefit if being developed according to upper ones [8].

Several comparisons of upper ontologies have been made for specific reasons, for example, when examining possiblities of upper ontology use in U.S. government and military [13] or for purpose of aligning heterogeneous ontologies automatically with intelligent software agents [9]. Comparisons are also usualy made when developing new upper ontology [5] or for theoretical contemplation purposes from the point of view of philosophy, where the term ontology originated [7].

According to continual research over years, upper ontologies can have valuable role in large number of research areas. There are many examples of their purpose and usability, especially considering integration of heterogeneous knowledge, ontology reuse and semantic interoperability. But idea of one unified upper ontology that could connect all others, both genaral and specific, is continuously present.
In [8] is claimed that many domain ontologies are not being developed according to upper ones or, if they are, such information isn't provided. For all ontologies to be integrated, it would be usefull that each of them has a connecting point to upper onotology. The question whether this connection can easily be detected intrigued us, so we decided to conduct an initial research. We chose two well-known upper ontologies and ten random domain ontologies to see whether and where are they connected. The purpose was to answer the following questions: Is there a potential for more thorough research? What course for further analysis can be suggested? Can more comprehensive research give a contribution (an additional considering factor, for exmple) to universal upper ontology development efforts?

2 SUMO and OpenCyc

Upper ontologies (top-level or foundation ontologies) "describe very general concepts that are common across the domains and give general notions under which all the terms in existing ontologies should be linked to" [3]. The term has also been used for broad general ontologies of a specific domain, for example, in life sciences [1], as a medium level between upper ontology and specific domain ontology. For this initial research we have chosen two upper ontologies that are continously beening developed over years and also have portals that allow users to browse terms online – Suggested Upper Merged Ontology (SUMO) and OpenCyc. If the future research is proven to be usefull, other well-known upper ontologies will also be included: Basic Formal Ontology – BFO (`http://www.ifomis.org/bfo`), General Formal Ontology – GFO (`http://www.onto-med.de/`

ontologies/gfo/), DOLCE (http://www.loa.istc.cnr.it/DOLCE.html), SOWA (http://www.jfsowa.com/ontology/index.html) and PROTON (http://www.ontotext.com/proton-ontology).

Suggested Upper Merged Ontology (SUMO)
http://www.ontologyportal.org
http://suo.ieee.org/SUO/SUMO/index.html

Developers of SUMO claim this upper ontology and its domain ontologies to be currently the largest formal public ontology and the only such ontology that has been mapped to all of the WordNet lexicon. Its development started in 2000 as a candidate for Standard Upper Ontology (SUO) and was presented to public in 2001, as ontology that "will provide definitions for general-purpose terms" and "will act as a foundation for more specific domain ontologies" [10]. It is written in first-order logic language called SUO-KIF (Standard Upper Ontology Knowledge Intercange Format) and translated to OWL. It is owned by IEEE and free to use.

SUMO is a modular ontology that consists of top-level (SUMO itself), Mid-Level Ontology (MILO) and other subontologies of various broad domains. Alltogether, it has about 25.000 terms and about 80.000 axioms, but SUMO itself is limited to 1.000 terms, 4.000 axioms and over 800 rules. Structure of this modular ontology is presented in Fig. 1.

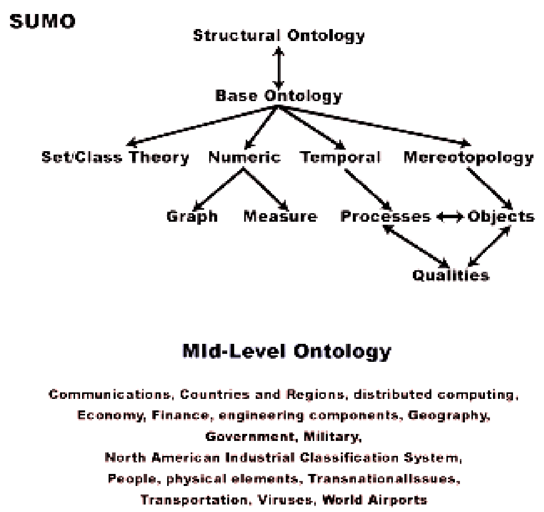

Figure 1. SUMO modular ontology

The highest (root) concept in SUMO is *Entity*. *Entity* has two subconcepts – *Physical* and *Abstract*. Concept *Physical* is further divided into *Object* and *Process*, while *Abstract* is divided into *Quantity*, *Atribute*, *Set or Class*, *Relation*, *Proposition*, *Graph* and *Graph Element*. Hierarchy then branches further to lower level concepts. Ontology content can be browsed online using system called Sigma Knowledge Engineering Environment (http://sigmakee.sourceforge.net/).

OpenCyc
http://opencyc.org
http://sw.opencyc.org
http:/www.cyc.com

OpenCyc is opensource of Cyc technology, also claiming to be largest and most complete general knowledge base today. Like SUMO, it is being considered by IEEE for a SUO standard. Development of Cyc began in 1984 and OpenCyc was released in 2002, as a result of improvement efforts described in OpenCyc white paper [11]. Cyc knowledge base is constantly growing and it is a formalized reperesentation of a large quantity of general human knowledge, "divided into thousands of 'microtheories' focused on a particular domain of knowledge, a particular level of detail, a particular interval in time, etc" [9]. It is written in language CycL and also translated to OWL.

Currently OpenCyc has over 239.000 terms and over 2.000.000 triples, as well as more than 69.000 links to external semantic data namespaces like DBpedia or WordNet. OpenCyc can be used for various purposes, for example: ontology development, expert systems, games and email prioritizing. In Fig. 2 is shown only small upper part of Cyc concept hierarchy, with highest (root) concept *Thing*. It can be seen that concepts in hierarchy are intertwined.

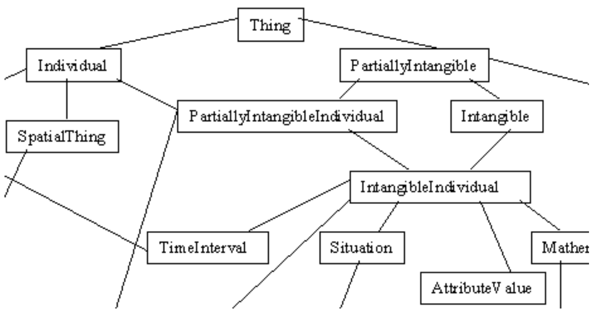

Figure 2. Upper part of CyC hierarchy

3 Connecting domain and upper ontologies

For the purpose of our research we have randomly chosen ten ontologies from various (broader and narrower) domains and explored where can they be connected with two chosen upper ontologies. On purpose, no specific criteria was used when selecting them, except that they all had to have the same knowledge representation language – Web Ontology Language (OWL). Those ontologies are (in alphabetical order):

– Black Pepper (http://www.dataindia.org/ontologies/OntoBlackPepper.owl);
– Cheminf (http://semanticchemistry.googlecode.com/svn/trunk/ontology/cheminf.owl);

- Colors (`http://kaiko.getalp.org/kaiko/ontology/colors.owl`);
- Family (`http://protege.cim3.net/file/pub/ontologies/family.swrl.owl/family.swrl.owl`);
- Finance (`http://www.fadyart.com/ontologies/data/Finance.owl`);
- GeoSkills (`http://i2geo.net/ontologies/current/GeoSkills.owl`);
- Monetary (`http://protegewiki.stanford.edu/images/d/de/Monetary_ontology_0.1d.zip`);
- PP (`http://sites.google.com/site/ppontology/PPOntology.zip`);
- Travel (`http://protege.cim3.net/file/pub/ontologies/travel/travel.owl`);
- Wine (`http://protege.cim3.net/file/pub/ontologies/wine/wine.owl`).

Table 1. Analysis of connections between domain and upper ontologies

Ontology Name	Ontology Description	Connection to SUMO	Connection to OpenCyc
Black Pepper	– ontology of diseases and insects that can attack black pepper – 3 root classes: Disease, Insects, Spices	– Disease – Disease Or Syndrome (Biological Attribute), 7 sub – Insects – Insect (Arthropod), 7 sub – Spices – Spice (Fruit Or Vegetable), 10 sub	– Disease – disease (type of ailment), 6 sub – Insects – insect (arthropod or vermin), 24 sub – Spices – seasoning (food ingredient or vegetable matter), 10 sub
Cheminf	– richly describes chemical entities with qualitative and quantitative attributes – 2 root classes: Entity, ObsoleteClass	– Entity – Entity, top concept, 2 sub – Obsolete (Class) – Subjective Assessment Attribute (Normative Attribute), 5 sub	– Entity – no match, top concept is thing – Obsolete (Class) – obsolete constant (CycL constant or FACtory irrelevant term), 1 sub
Colors	– small ontology that describes specific individuals of colors – 1 root class: color	– Color – Subjective Assessment Attribute (Normative Attribute), 5 sub	– Color – color (color perception category or type of object or visible feature), 9 sub
Family	– ontology that shows relationships among family members – 3 root classes: entity, gender, person	– Entity – Entity, top concept, 2 sub – Gender – Sex Attribute (Biological Attribute), no sub – Person – Human (Cognitive Agent or Hominid), 5 sub	– Entity – no match, top concept is thing – Gender – gender of person (person) or gender (living thing), no sub – Person – person (human, legal agent, sentient, social agent or thing existing stably in time that is not an organization), several hundred sub
Finance	– ontology about financial issues in securities handling – 10 root classes: Account, Contact, FinancialInstrument, Organization, Party, ProcessAndProcedure, Risk, Service, Temporal, Value Partition	– Account – Social Interaction (Intentional Process), 15 sub – Contact – unique Identifier (instance), x sub – Financial Instrument – Financial Instrument (Certificate), 9 sub – Organization – Organization (Agent or Group), no sub – Party – Agent (individual), x sub – Process and Procedure – Process (Physical), 11 subclasses; Procedure (Proposition), 6 sub – Risk - Subjective Assessment Attribute (Normative Attribute), 5 sub – Service – Cooperation (Social Interaction), no sub – Temporal – Time Measure (Constant Quantity), 2 sub – (Value) Partition – Separating (Dual Object Process), 5 sub	– Account – account (design), 2 subclasses – Contact – contact information (pit), no sub – Financial Instrument – financial instrument (authorized agreement, financial asset of investment vehicle), 10 sub – Organization – organization (agentive artifact, group of intelligent gents acting together, group of things, intelligent agent,, social agent, structure), several hundred sub – Party – generic agent (agent, partially intangible individual or thing existing stably in time), several hundred sub – Process and Procedure – processing (action on an object), 10 sub; procedure-specification (protocol), no sub – Risk – risking (intentional action), 4 sub – Service – service (helping), more than hundred sub – Temporal – time (non aspectual quantity, non-negative quantity or number, one dimensional quantity or physical measurable quantity), 4 sub – (Value) Partition – division (separating), 1 sub
GeoSkills	– ontology about mathematical competences, topics and educational level for European schools – 9 root classes: Competency, EducationalLevel, EductionalPathway, EducationalProgram, EducationalReligion, NamableBit, Resource, Topic, Class	– Competency - Subjective Assessment Attribute (Normative Attribute), 5 sub – Educational Level - Educational Organization (Organization), 3 sub – Educational Pathway - Educational Organization (Organization), 3 sub – Educational Program – Normative Attribute (Relational Attribute), 2 sub – (Educational) Region – Region (Object), 2 sub – (Namable) Bit - Subjective Assessment Attribute (Normative Attribute), 5 sub – Resource - Subjective Assessment Attribute (Normative Attribute), 5 sub – Topic – Proposition (Abstract), 17 sub	– Competency – competence (manner in which an action is performed), no sub – Educational Level – educational organization (organization), 11 sub – Educational Pathway – educational organization (organization), 11 sub – Educational Program – curriculum (design), no sub – (Educational) Region – geographical region (geopolitical entity or region, place that is also a geographical thing, tangible thing or underspecified surface), more than hundred sub – (Namable) Bit – part (section, tangible thing), 45 sub

		– Class - Subjective Assessment Attribute (Normative Attribute), 5 sub	– Resource – resource (thing that exists in time), 3 sub – Topic – topic (conventional classification type or facet instance collection), 4 sub – Class – class (meeting or organized event), no sub
Monetary	– simple ontology about currencies – 11 root classes: Actor, Agreement, Denomination, Issuance, Minting, Policy, Role, Symbol, Trade, Value, ValuePartition	– Actor – Human (cognitive agent or hominid), 5 sub – Agreement – Agreement (proposition), no sub – Denomination - Class (Set or class), 1 sub – Issuance – Making (Intentional Process), 7 sub – Minting – Manufacture (Making), 2 sub – Policy – Argument (Proposition), 3 sub – Role – holds Obligation (instance), x sub – Symbol – Character (Symbolic String), 2 sub – Trade – Financial Transaction (Transaction), 24 sub – Value – monetary Value (instance), x sub – (Value) Partition – Separating (Dual Object Process), 5 sub	– Actor – actor (event or thing existing stably in time), no sub – Agreement – agreement (policy), 29 sub – Denomination – naming (intentional action or specifying), 1 sub – Issuance – emitting (transfer event with well-defined from-location or translocation), 6 sub – Minting – minting (intelligent agent activity or making something), no sub – Policy – policy (design), 6 sub – Role – role (extensional representation predicate), 2 sub – Symbol – symbol (thing), no sub – Trade – financial instrument (authorized agreement, financial asset of investment vehicle), 10 sub – Value – monetary value (economical quantity, measure of utility or one dimensional quantity), 29 sub – (Value) Partition – division (separating), 1 sub
PP	– barley plant protection ontology – 10 root classes: Abnormality, Case, Cultural_Practice, Disorder, Environmental_Condition, Growth_Stage, Material, Observation, Organism, Plant_Parts	– Abnormality – Subjective Assessment Attribute (Normative Attribute), 5 sub – Case – Subjective Assessment Attribute (Normative Attribute), 5 sub – Cultural Practice – Agriculture (Maintaining), 1 sub – Disorder – Disease Or Syndrome (Biological Attribute), 7 sub – Environmental Condition – Attribute (Abstract), 3 sub – Growth Stage – Growth (Autonomic Process), no sub – Material – Substance (Self Connected Object), 24 sub – Observation – Measuring (Calculating), no sub – Organism – Organism (Agent or Organic Object), 7 sub – Plant Parts – Plant Anatomical Structure (Anatomical Structure), 9 sub	– Abnormality – abnormality (social quantity), no sub – Case – event (landmark-unspecified), situation), several hundred sub – Cultural Practice – agriculture (human activity, purposeful composite physical and mental activity or unnatural thing), 5 sub – Disorder – type of ailment (KE clarifying collection type, type of object or type of temporally stuff-like thing), several hundred sub – Environmental Condition – sustainable agriculture (agriculture), no sub – Growth Stage – biological growth event (biological event, physical creation event or physical growth), 13 sub – Material – tangible thing (agent, location, thing existing stably in time, three dimensional thing or unspecified surface), several hundred sub – Observation – observation (remark), no sub – Organism – organism (living thing, spatially continuous thing), 50 sub – Plant Parts – plant part (organism part or plant or part of a plant), 43 sub
Travel	– example ontology for tutorial purposes – 5 root classes: Accommodation, AccomodationRating, Activity, Contact, Destination	– Accommodation – Traveler Accommodation (attribute), x sub – (Accommodation) Rating – Rating Attribute (Subjective Assessment Attribute), 6 sub – Activity – Intentional Process (Process), 19 sub – Contact – unique Identifier (instance), x sub – Destination – destination (instance), x sub	– Accommodation – lodging (organization with individual clients or service organization), 3 sub – (Accommodation) Rating – rating type criterion (evaluative quantity), no sub – Activity – action (event or expression), more than hundred sub – Contact – contact information (pit), no sub – Destination – destination (translocation, geographical thing or location), no sub
Wine	- demonstrational ontology about wines - 9 root classes: Consumable Thing, Fruit, NonConsumable Thing, Region, Vintage, Vintage Year, Wine, WineDescriptor, Winery	– Consumable Thing – capability (instance), x sub – Fruit – Fruit Or Vegetable (Plant Anatomical Structure or Reproductive Body), 17 sub – Nonconsumable Thing – x (Entity), all sub – Region – Region (Object), 20 sub – Vintage – age (instance), x sub – (Vintage) Year – Year (Time Interval), 2 sub – Wine – Wine (Alcoholic Beverage), 1 sub – (Wine) Descriptor – Attribute (Abstract), 3 sub – Winery – Stationary Artifact (Artifact) , 23 sub	– Consumable Thing – consumable product (goods), 30 sub – Fruit – fruit (external anatomical part, plant part, portable object, rigid portable object or solid object), 44 sub – Nonconsumable Thing – x (thing), all sub – Region – region (place), no sub – Vintage – vintage thing (artifact or consumer durable), no sub – (Vintage) Year – year (date), 9 sub – Wine Descriptor – attribute (abstraction), 1 sub – Winery – wine maker (food product company), no sub

Our condition for a domain ontology to be connected with upper one was that its root concepts actually represent further branching of concepts of an upper ontology. Therefore, for each root concept (class) in ten chosen ontologies analysis was made to establish:

– if exact match can not be found, which concept in upper ontology can be considered as a substitute for concrete root concept from domain ontology;
– which is a direct upper ontology superclass (or superclasses) of a root domain ontology class;
– how many subclasses root class match/substitute has in upper ontology.

Analysis is shown in Table 1 where entry for each root class is written as follows: *RootClassName – RootClassMatch/Substitute (Direct Superclass-es), Number of subclasses*. If *RootClassMatch/Substitute* is of type instance or attribute, then it can not have subclasses, which is indicated with "x sub". It should be taken into consideration that number of subclasses depends on domain and other specific ontologies whose concepts are also stored in knowledge base of upper ontologies SUMO and OpenCyc.

As it can be seen, we have included in initial research only classes and not properties from domain ontologies. In total, there were 63 root classes in 10 ontologies and only 3 semantically repeated (entity, contact, partition value) – in total 60 different classes. From descriptions in Table 1 it is obvious that SUMO and OpenCyc to some extent have different naming conventions as well as structure and organization. If several substitutes for a certain class were possible, we made selection according to their semantic similarity to the meaning of a root class in its domain ontology. It is of significance that we found exact match (strictly the same word, only singular/plural is accepted) only for 10 classes in SUMO and 24 in OpenCyc. If concepts with additional word(s) are accepted (for example, if according to domain ontology semantics "monetary value" can be match for "value"), then there are 15 matches in SUMO and 29 in OpenCyc. With addition of similar words (for example, "risk" and "risking") there are few more matches – 32, but only in OpenCyc.

4 Conclusions and further work

Initial research about connections between domain and upper ontologies has been conducted and it showed that, even though exact match or substitute can be found in upper ontologies for each root class, relationships are not in all cases simple to locate. And properties as well as better part of existing upper ontologies were not even included into analysis.

As expected, we have learned that a large number of ontologies is not developed exactly according to upper ontologies. We tested only two of them, but they are the most general ones. It became obvious that for the purpose of establishing connections between domain and upper ontologies, a specific set of concept mappings must be made.

Our initial research showed that only for 16,7% (SUMO) to 24% (OpenCyc) of domain root classes exact match can be found in upper ontologies. If additional and semantically similar words are allowed, then a matching percentage is from 25% (SUMO) to 53,3% (OpenCyc). Therefore, for at least half of classes there is no exact match and, as can be seen in Table 1, sometimes concepts on more general level should be used as substitutes.

Although this was only initial research with small number of domain ontologies, according to results we can conclude that there is a potential for more thorough research, where other upper ontologies and domain ontology properties would be included. To obtain more significant results, more extensive research with more explicitly defined factors should be conducted, some of them being:

– strict match and loose match ratio for classes and properties;
– difference in root and substitute concept generalization level;
– use of match/substitute subclasses in domain ontology subclasses.

With even more comprehensive research and proposal for a set of concept mappings that would be used for connecting all domains with upper ontologies, a modest contribution can also be made in universal upper ontology development efforts.

References

[1] Beißwanger E, Schulz S, Stenzhorn H, Hahn U: BioTop: An Upper Domain Ontology for the Life Sciences, Applied Ontology, Vol. 3, Iss. 4, IOS Press, Amsterdam, Netherlands, 2008, pp. 205-212

[2] Colomb R M: **Use of Upper Ontologies for Interoperation of Information Systems**, A Tutorial, Technical Report 20/02 ISIB-CNR, 2002, available at `http://www.loa.istc.cnr.it/Papers/ISIB-CNR-TR-20-02.pdf`, Accessed: 20th April 2012.

[3] Gómez Pérez A, Fernández-López M, Corcho O: **Ontological Engineering**, Springer, Berlin, Germany, 2004.

[4] Guarino N: **Some Organizing Principles For a Unified Top-Level Ontology**, AAAI Technical Report SS-97-06, 1997, available at `http://www.aaai.org/Papers/Symposia/Spring/1997/SS-97-06/SS97-06-008.pdf`, Accessed: 21st April 2012.

[5] Herre H, Heller B, Burek P, Hoehendorf R, Loebe F, Michalek H: **General Formal**

Ontology (GFO), Part I: Basic Principles, Onto-Med Report no. 8, 2006, available at `http://www.onto-med.de/publications/2006/herre-h-2006-a.pdf`, Accessed: 16th April 2012.

[6] Jain P, Yeh P Z, Verma K, Vasquez R, Damova M, Hitzler P, Sheth A: **Contextual Ontology Alignment of LOD With an Upper Ontology: A Case Study With Proton**, Proceedings of 8th Extended Semantic Web Conference, Part 1, May 29 – June 2 , Heraklion, Greece, 2011, pp.80-92.

[7] Jansen, L: **Categories: The Top-Level Ontology**, In (Munn K, Smith B (eds.)): Applied Ontology, An Introduction, ontos verlag, Hausenstamm bei Frankfurt, Germany, 2009, pp. 173-196.

[8] Keet C M: The **Use of Foundational Ontologies in Ontology Development: An Empirical Assessment**, Proceedings of 8th Extended Semantic Web Conference, Part 1, May 29 – June 2 , Heraklion, Greece, 2011, pp.321-335.

[9] Mascardi V, Cordi V, Rosso P: **A Comparison of Upper Ontologies**, Technical Report DIS-TR-06-21, 2006, available at `http://www.disi.unige.it/person/MascardiV/Download/DISI-TR-06-21.pdf`, Accessed: 15th April 2012.

[10] Niles I, Pease A: **Towards a Standard Upper Ontology**, Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, October 17-19, Ogunquit, Maine, USA, 2001, pp. 2-9.

[11] Reed, S.L., Lenat D.B.: **Mapping Ontologies into Cyc**, AAAI 2002 Conference Workshop on Ontologies and the Semantic Web, July 29, Edmonton, Canada, 2002, available at: `http://www.cyc.com/doc/white_papers/mapping-ontologies-into-cyc_v31.pdf`, Accessed: 4th April 2012.

[12] Sanchez-Alonso S, Garcia-Barriocanal E: **Making Use of Upper Ontologies to Foster Interoperability Between SKOS Concept Schemes**, Online Information Review, Vol. 30, Iss. 3, Emerald Group Publishing Limited, Bingley, UK, 2006, pp.263 - 277

[13] Semy S K, Pulvermacher M K, Obrst L J: **Toward the Use of an Upper Ontology for U.S. Government and U.S. Military Domains: An Evaluation**, MITRE Technical Report 04B0000063, 2004, `http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA459575`, Accessed: 16th April 2012.